Detecting the True Extent of Introgression during Anthropogenic Hybridization

Detecting the true extent of introgression during anthropogenic hybridization

S. Eryn McFarlane[1,2]* & Josephine M. Pemberton[1]

1. Institute of Evolutionary Biology, School of Biological Sciences, University of Edinburgh, Charlotte Auerbach Road, Edinburgh, United Kingdom, EH9 3FL
2. Department of Biology, Lund University, Sölvegatan 37, 223 62 Lund, Sweden
*author for correspondence: eryn.mcfarlane@ed.ac.uk

Abstract

Hybridization among naturally separate taxa is increasing due to human impact, and can result in taxon loss. Previous classification of anthropogenic hybridization has largely ignored the case of bimodal hybrid zones, in which hybrids commonly mate with parental species resulting in many backcrossed individuals with a small proportion of introgressed genome. Genetic markers can be used to detect such hybrids, but until recently too few markers have been used to detect the true extent of introgression. Recent studies of wolves and trout have used thousands of markers to reveal previously undetectable backcrosses. This improved resolution will lead to increased detection of late generation backcrosses, shed light on the consequences of anthropogenic hybridization, and pose new management issues for conservation scientists.

Anthropogenic hybridization

Anthropogenic hybridization (see Glossary), in which human disturbance leads to range overlap and **hybridization** of previously reproductively isolated populations or species is a growing conservation concern [1-3]. With increased human-generated movement of species into new ranges, there is an increasing number of cases of hybridization between species that were historically **allopatric** [4]. Disturbance of habitats can also result in a breakdown of reproductive isolation between previously isolated, **sympatric** species [1]. **Introgression** is usually hard to detect from phenotypes and there is growing evidence that backcrossing has often proceeded further than is detectable by low density genetic marker panels. In this article we make the case that genomic approaches are essential and increasingly available to disentangle late generation backcrosses from parental populations after introgression has occurred.

The benefits of anthropogenic hybridization

There are possible benefits of anthropogenic hybridization. Policy makers can use hybridization as a management tool to help endangered populations. In 'genetic rescue' programs (i.e. breeding programs designed to release small populations from inbreeding depression), individuals from a closely related population or subspecies are introduced to an inbred population to manage inbreeding depression. For example, when Florida panthers (*Puma concolor coryi*) were threatened due to inbreeding depression, eight Texas panthers (*P. concolor cougaur*) were introduced. The **hybrid** kittens survived better, and the population is now recovering [5]. Approximately 90% of such genetic rescue attempts have been successful, showing that anthropogenic hybridization is a viable conservation method [6]. Adaptive introgression ('evolutionary rescue') in which beneficial alleles from an introduced population are selected for in hybrid individuals is another possible benefit of anthropogenic hybridization. For example, a segment of chromosome 15 that has naturally introgressed from *Populus balsamifera* into *P. trichocarpa* appears to allow *P. trichocarpa* to live in colder, drier areas than *P. trichocarpa* individuals without this haplotype [7]. This suggests that there is potential for adaptive introgression to facilitate

evolutionary rescue of populations at risk of extinction due to climate change [8], although such genomic management of at risk populations much enabling research, and should be approached with caution [9, 10].

The problems with anthropogenic hybridization

Anthropogenic hybridization can cause problems for native species. When no offspring or sterile offspring are produced, reproductive effort is wasted [11]. When fertile F1s are formed, introgression between the two previously diverged species is possible. There are two reasons why even low levels of introgression of non-native alleles are of concern from a conservation perspective. First, if all individuals of a species are hybrids then the species as it was is extinct. This has been termed 'extinction by hybridization' [11-15]. Note, however, that there may still be many copies of the native alleles represented in the population, so long as the population itself is large enough, and from a 'gene view point' we may be content with this mode of conservation [16].

The second problem with hybridization is that introgression and recombination break up linked gene complexes, and non-native alleles that are favoured (or no longer in linkage with deleterious alleles) can be swept to fixation [17]. While this leads to an initial increase in biodiversity (because alleles from both the native and non-native populations are present) as non-native alleles sweep to fixation, native alleles are lost. If we again take a gene view point of biodiversity, any alleles lost from the native population are a loss in biodiversity from the system. For example, non-native alleles at three out of 68 genetic markers have gone to fixation in some populations of California Tiger Salamanders (*Ambystoma californiense*) after hybridization with Barred Tiger Salamanders (*A. mavortium*) [18]. This has occurred in California Tiger Salamander populations that are nearly 100km from the original Barred Tiger Salamander introduction site, suggesting that these alleles have higher fitness than the native, California Tiger Salamander alleles that they have replaced [18].

Goals of studies of anthropogenic hybridization

Studies of anthropogenic hybridization have different goals. A researcher might be interested to know if hybridization has occurred at all in a population to determine whether it should provide the breeding stock for new populations, and or whether it should be quarantined because of hybridization. Relatively few informative markers are needed to detect individuals of hybrid origin in any particular population, as the detection of any non-native allele is a clear indication of hybridization [19].

However, if a researcher wishes to understand more about the underlying process of hybridization and introgression, then many more markers are required. Specific goals might include: to select individuals for breeding programs; to understand the relationship between genotype and phenotype; to understand the type of hybrid system involved (see next section); and to investigate mating patterns and fitness. For any of these goals, it is ideal to quantify individual **admixture** accurately, and to do this this hundreds or thousands of informative markers may be required (see below).

Classifying hybridization

To assist researchers and policy makers in addressing anthropogenic hybridization, Allendorf and colleagues [11] categorized hybridization outcomes. Types 1-3 applied to naturally-occurring hybridization while Types 4-6 applied to anthropogenic hybridization. Type 4 results in few or sterile F1 hybrids, and is characterized by wasted reproductive effort. Type 5 results in a **hybrid swarm** with widespread introgression into particular populations, but some populations do not experience hybridization at all. Finally, Type 6 results in a complete hybrid swarm following break down of reproductive isolation between species across all populations [11].

Three axes of variation determine the outcome of anthropogenic hybridization: differences in hybrid fitness, time since **secondary contact**, and mating patterns of hybrids. Time since secondary contact and mating patterns of hybrids were not explicitly considered in Allendorf et al's original categorization. Type 4 differs from Types 5 and 6 along an axis of hybrid fitness, where intrinsic post

zygotic isolation affects hybrids in Type 4, but not in Types 5 or 6. This results in little to no backcrossing in Type 4 hybrid zones, as hybrids are extremely unfit compared to parental species. This decrease in hybrid fitness must be extreme, as even with a 90% decrease in fitness, the proportion of hybrids in a hybridizing population is expected to increase [20].

We suggest that the only difference between Allendorf et al's [11] Type 5 and Type 6 is time since secondary contact. When an F1 reproduces, all of its offspring and descendants are admixed to some extent [20]. If Type 5 characterizes a system where only one or few populations have introgression, Type 6 is the logical outcome of this same system, assuming random mating and sufficient time for migration between populations. Thus, we consider Type 5 and Type 6 to be the same, both hybrid swarms with a breakdown of assortative mating, in which hybrids have the same mating success as either of the parental species individuals, and common enough that hybrid x hybrid matings occur.

When there is a preference among hybrids for parental species phenotypes, or hybrids are very rare, we expect a different pattern of introgression. Backcrossing into the parental species leads to an increasingly large number of individuals with a small proportion (<10%) of their genome that is from the opposite species. As backcrossing continues, morphological differences between parental species and backcrossed individuals lessen, making it more and more difficult to detect a backcross using only phenotypic traits. This results in many hybrid individuals with very small proportions of another genome, although with a maintained bi-modal distribution of trait values between the two parental species (Figure 1). From a conservation perspective we consider this to be a worst-case scenario as these introgressed individuals are very difficult to detect. This can be contrasted with a general lack of assortative mating, in which hybrid individuals are as likely to breed with other hybrid individuals as with parental species (leading to a hybrid swarm), or, in the unlikely event of true assortative mating, where hybrid individuals preferentially breed with each other, which would lead to the eventual formation of a hybrid species e.g. [21]. The contrast between hybrid zones with unimodal distributions of traits and admixture

scores and those with bimodal distributions has previously been described in the context of naturally occurring hybrid zones [22], but does not yet seem to inform studies of anthropogenic hybridization.

The distribution of hybrid scores in a system at equilibrium varies depending on ecological factors that can affect hybrid fitness, and hybrid encounter rate. Extrinsic post zygotic isolation can vary according to ecological factors, affecting the ability of hybrids to successfully mate and reproduce [23]. Further, stochastic factors, particularly when hybrids are rare, or management might alter the reproductive success of hybrid individuals in wild systems. However, if hybrids are fertile, the proportion of hybrid individuals in all populations should increase [20], leading to the extreme end points of majority hybrid populations which either follow a hybrid swarm or **bimodal hybrid zone** distribution.

Key considerations for genetic analyses of anthropogenic hybridization

Published studies of anthropogenic hybridization generally follow a similar protocol. Researchers use codominant marker genotypes to estimate divergence between the two species [24] and then use a clustering approach such as STRUCTURE [25-28], or ADMIXTURE [29, 30] to partition individuals into different genetic groups (K). Those individuals with an admixture score (Q) intermediate to the extreme admixture scores associated with parental species individuals are designated hybrids. Many studies then use HYBRIDLAB [31, 32] or similar methods to simulate hybrid genotypes from the sampled genotypes to assess the **efficiency** (i.e. type II error rate, rate of assigning hybrid individuals as parental species), and **accuracy** (i.e. type I error rate, rate of erroneously assigning parental species individuals as hybrids; [33]). The 'overall performance' of an analysis is the product of efficiency and accuracy and this performance can be used to assess the reliability of the study itself [33]. Here we outline some best practices and points to consider in order to avoid underestimation of the extent of hybridization.

Divergence between parental species

It is highly relevant to have an estimate of divergence between the focal species in the absence of hybridization. **F_{st}** is often reported in studies of anthropogenic hybridization, but is rarely used to motivate the marker density deployed for estimates of individual admixture, typically because the same markers are used to determine both F_{st} and individual Q estimates. Simulations have clearly shown that species (or subspecies) with lower divergence will require more markers to accurately estimate admixture, because of shared polymorphisms between them, leading to fewer **diagnostic markers** [33]. While it might not be practical to use markers to estimate F_{st} and then determine how many markers are needed to estimate individual admixture scores, an initial assessment of F_{st} will hint at how much power a system has to detect advanced backcrosses.

Historical admixture

Many systems have a history of repeated secondary contact and hybridization. Documenting historical admixture using genomic resources can determine whether the introgression found is due to recent, anthropogenic forces, or to natural causes, which will change the conservation status of the situation [34, 35]. There are techniques for detecting historical admixture. For example, the ABBA-BABA test can be used to determine if there has been historical introgression from a third species or population into each of two closely related sister taxa, to explain variation that is not well explained by a null assumption of bifurcating phylogeny [36]. This technique can be applied to either sequences of single individuals from each population, or to multiple individuals from each population [37], and can be used to indicate historical (hundreds to thousands of generations before present) admixture. Similarly, *δaδi* analyses can be used to determine how well different demographic models fit the pattern of variation in the data, where demographic models can include admixture at different time points [38]. For example demographic modeling was used to demonstrate that hybridization between golden-winged (*Vermivora chrysoptera*) and blue winged warblers (*V. cyanoptera*) has probably been occurring since the original species split, and is not solely due to anthropogenic forces [39]. Finally, researchers can

examine the length of haplotype blocks that are identical by descent, as linkage disequilibrium decays over time due to recombination [40, 41]. The distribution of haplotype block lengths should follow a Poisson distribution [41] and deviation from this distribution can be used to infer population admixture over both short (tens of generations) [42] and long time spans [41]. These and other techniques for disentangling historical and contemporary admixture are reviewed in [43].

Generations since secondary contact and recombination rates

It is important to estimate the number of generations since secondary contact to estimate the potential number of backcross generations in a system. This estimate might have substantial uncertainty, but in many cases of anthropogenic hybridization there are historical records that suggest when a non-native species was first introduced or sighted that can be combined with typical generation times for the taxa involved. The expected proportion of invasive genome in a backcrossed individual halves with each successive generation of backcrossing [44].

Recombination each generation leads to less linkage disequilibrium between non-native loci, which means that genotype at a species-specific marker in one position is less informative about surrounding, un-sampled loci. For example, genomic regions with high recombination rates were found to be associated with more introgression of the non-native genome in replicate swordtail (*Xiphophorus birchmanni and X. malinche*) hybrid zones [17]. Due to obligatory crossing over, which is expected to occur once per chromosome arm [45], at least twice as many markers as there are chromosome arms are needed to cover each independent section of the genome. In some cases, there is a species-specific estimate of recombination (e.g. [46]), or one can refer to taxon-specific patterns. For example, there is as much as 10 times more recombination in avian genomes than in mammalian genomes [47]. Additionally, information on recombination rate can be combined with genomic methods examining haplotype block lengths to date introgression events (as discussed above). We discuss how many markers are needed further in Box 1.

252

253 **Assessing the power of markers**

254 Many studies of anthropogenic hybridization assess the power of genetic
255 markers used by simulating hybrid genotypes and then determining the power
256 the markers have to detect these hybrid genotypes [48]. When assessing the
257 power of markers in this way, it is important to ensure that the biology of the
258 system is reflected in the simulation. In particular, if the two species of interest
259 have been in contact for many generations and F1s are thought to be fertile
260 (Figure 1), then simulations should account for the possibility of many
261 generations of backcrossing. This is rarely done in conservation genetic studies -
262 many studies simulate backcrosses to assess the power of their markers, and
263 find low power to detect even first generation backcrosses, for example finding
264 less than 80% of first generation backcrosses are properly assigned [49, 50].
265 Further information obtained from laboratory or field studies, such as
266 asymmetry in hybrid fertility (e.g. between sexes, Haldane's Rule [51] or
267 according to the species of the mother of the F1, Darwin's Corollary [52]), should
268 also be included in simulations. For example, if previous laboratory work has
269 established that backcrossing is largely unidirectional because of decreased
270 fitness of hybrid individuals in the opposite direction (as expected by Darwin's
271 Corollary) or due to the relative abundance of the parental species, then
272 mitochondrial markers should be integrated into future analyses to add power to
273 detect hybrids.

274

275 **Defining hybrid individuals**

276 To be defined as a hybrid, a focal individual must be genetically differentiated
277 from both parental species. Parental species are assumed to have an admixture
278 (Q) score of 0 or 1, although because of error (e.g. non-diagnostic markers,
279 genotyping errors), very few individuals will have an estimated score of exactly 0
280 or 1. Any score in between indicates a hybrid [25]. It is typical for a researcher to
281 set a Q score as a cut-off for hybrid individuals, so any individual above (or
282 below) this score is considered parental. Thresholds are determined either by
283 power, specifically, at what level can the markers differentiate between hybrids
284 and parental species, or by the number of acceptably mis-matched markers, e.g.

one allele indicative of the other species might be an error, but two markers suggest hybridization [53]. These thresholds can range widely between studies, from 0.8 [54] to 0.999 [30] in relation to a parental species score of 1.0. Determination of the threshold is a balancing act between Type I and Type II errors, in which the researcher must decide whether it is better to mistakenly assign a parental species individual as a hybrid (Type I; too low 'accuracy'; [33]) or assign hybrid individuals as parental types (Type II; too low 'efficiency'; [33]). If the researcher accepts a higher level of Type II errors, they consider advanced backcrosses as parental species. For example, an admixture score threshold of 0.8 would include most second-generation backcrosses (87% of the genome is species A, 13% of the genome is species B on average) as parental species. Similarly, with a Q score of 0.9, third-generation backcrosses (average of 93% species A) would be included as parental species individuals.

There are two ways to ameliorate error introduced in species assignment using thresholds. One obvious way is to employ more markers (Box 1), which increases the power of a study and allows the setting of thresholds approaching 0 and 1. Studies that have used thousands of markers use the most stringent thresholds e.g. [30]. A second solution to the threshold problem is to do away with them entirely. Rather than assigning individuals to species classes based on point estimates, it is more appropriate to use **credible** or **confidence intervals** around point estimates which capture uncertainty in the marker system appropriately (Box 2). In this scenario any individual with a credible interval overlapping 0 or 1 is considered a parental species and all others are considered hybrids.

An additional problem in separating hybrid individuals from parental species is that some hybrids, particularly later generation of backcrosses, will be homozygous for all sampled diagnostic loci by chance. This is due to increased variation around the proportion of genome inherited from each parental species with each generation of backcrossing ([44]; Box 1). The hybrid nature of these individuals will be undetectable, and they will be classified as parental species, even though unmarked genome regions may be introgressed. Increasing the

number of markers increases the probability of sampling a hybrid individual at loci that are heterozygous or homozygous for alleles representative of both parental species (Box 1).

Higher density markers to identify bimodal hybrid zones

When researchers apply higher density marker panels to examples of anthropogenic hybridization, they generally uncover more backcrossed individuals compared to studies using low-density panels, and can draw more accurate conclusions about the system. These newly-detected backcrosses are often genetically very similar to the parental species, with less than 10% introgression, indicative of a bimodal hybrid zone. For example, in a study of Italian wolves that hybridize with domestic dogs, use of 170,000 SNPs found that hybridization had occurred 3 -5 generations prior to sampling [30]. This multi-generation backcrossing was not detectable in the population when 18 microsatellite markers were used [49]. Further, while very few individuals were found to have Q scores between 0.25 and 0.75, as would be expected in a hybrid swarm with a complete breakdown of reproductive isolation, 62% of sampled Eurasian wolves had a small proportion (<5%) of admixture with domestic dogs [55]. The Eurasian wolf – domestic dog system has the distribution of admixture scores and phenotypes that characterizes a bimodal hybrid zone with some degree of mating preference for parental phenotypes, or rare intermediate hybrids. In this system, most individuals are either phenotypically dog-like with extreme Q scores at one end of the distribution, or phenotypically wolf-like with Q scores at the other end of the distribution. There are few individuals with intermediate Q scores and phenotypes. This can be contrasted with the westslope cutthroat (*Oncorhynchus clarki lewisi*) – rainbow trout (*O. mykiss*) system, which has also recently been genotyped using 3180 diagnostic SNPs [56]. While the increase in number of markers did lead to increased detection of advanced backcrosses, there were also many individuals with intermediate Q scores and phenotypes [56, 57]. This suggests that the westslope cutthroat-rainbow trout system is a hybrid swarm that has little assortative mating.

Designing an ideal study of an anthropogenic hybrid zone

When embarking on a study of anthropogenic hybridization, there are many considerations in deciding on the genetic resources to be used (Box 1). As whole genome sequencing (WGS) becomes cheaper [58], conservation biologists should consider whether WGS is the best way forward. Firstly, WGS data allows for detection of heterogeneity of introgression across the genome. If conservation biologists truly adopt a 'gene view point' of hybridization [16] then individuals ought to be classified based on whether they carry specific alleles at identified loci, rather than by overall Q scores (but see [10] for a discussion of the difficulty of implementing this approach). Secondly, WGS enables the researcher to distinguish between historical and contemporary introgression. Finally, we anticipate that the use of WGS will result in more diagnostic or **ancestry informative markers** being detected, and thus make studies more powerful. Researchers will be more confident in their estimates of individual admixture, and will report the power and confidence associated with their analyses (Box 2). While the bioinformatics skills required to assemble a genome and call SNPs may seem intimidating, we believe that 1) these are skills are now routinely taught in universities and 2) WGS presents an additional opportunity for conservation biologists to collaborate with speciation geneticists (Box 3). Another consideration is that high quality DNA is needed for the most accurate assemblies, although progress is being made towards high quality sequences from poor quality samples (e.g. [59]). While the use of WGS is more expensive than microsatellite marker studies, when the cost of microsatellite markers, including the cost of labour, was compared to the use of SNP markers in European wolves, SNPs were less expensive if at least 24 samples were genotyped [60]. This suggests that the use of thousands of variable genome wide markers (e.g. from ddRAD [61]) may represent a practical middle ground for conservation biologists, depending on the history and biology of the system. Taken together, we believe that the best way forward to accurately detect backcrossing in studies of anthropogenic hybrid zones is to routinely use higher density markers, including WGS when possible.

Concluding Remarks

Advanced backcrosses are unlikely to have been detected with many of the methods that biologists studying anthropogenic hybridization have used to date. Most studies of anthropogenic hybridization have used fewer than 20 markers [13], too few to reliably detect individuals that are more than two generations backcrossed [33], unless markers are perfectly species diagnostic [44]. For this reason, it is rare for studies to consider backcrossed individuals past the second generation of backcrossing, regardless of the number of generations that have passed since secondary contact. Here, we suggest that studies should attempt to go much further. By accounting for the number of generations since secondary contact and increasing the density of genetic markers accordingly, many more backcrossed individuals will become distinguishable from the parental populations. We echo the call for more genetic markers to be used in these studies to allow for higher accuracy and efficiency [1, 3, 13, 33, 62], particularly since we have now entered the genomics era, making tens or hundreds of thousands of markers obtainable even in non-model systems [58]. It seems likely that anthropogenic hybridization will only increase in frequency and result in increased gene flow between previously isolated species [1]. The increase in number of markers and associated power will also open up the opportunity to ask new questions in these systems, parallel to those speciation biologists explore in natural hybrid zones (Box 3). There are new challenges with increased marker density, but a genomic approach to studying these systems will help researchers to detect backcrosses and make the best policy recommendations.

Acknowledgements: We would like to thank Konrad Lohse and two anonymous reviewers for valuable feedback on this manuscript. S.E.M. was funded by the Swedish Research Council (Vetenskapsrådet, International Postdoc 2017-00499).

Additional Elements

Glossary

- **Anthropogenic hybridization:** the breakdown of reproductive isolation between two species due to human action, including but not limited to, species introduction, habitat disturbance or escape of domestic species.
- **Accuracy:** the proportion of identified hybrids that are actually of hybrid ancestry [33]. A low accuracy suggests a high rate of type I errors, in which parental species individuals are erroneously assigned as hybrids.
- **Admixture**: the mixing of genomes from structured or diverged populations
- **Allopatry**: species in non-overlapping ranges
- **Ancestry informative markers:** genetic markers with substantial allele frequency differences across populations, which can be used to assign individuals to each population [63]
- **Bimodal hybrid zone:** a hybridizing population in which preference for parental phenotypes, or scarcity of hybrids with which to mate, results in a population that includes few F1 hybrids, and many backcrossed individuals with a low level of introgression that often resemble the parental species in phenotype. Can be unimodal (if backcrossing is into just one parental species) or bimodal (backcrossing into both parental species) [22]
- **Credible interval:** the range of possible values surrounding a point estimate, representing the uncertainty in the estimate
- **Diagnostic markers:** markers with fixed allele differences across populations
- **D_{xy}:** an absolute measure of genetic differentiation, calculated as the proportion of nucleotides that differ between two homologous sequences within the same or different population.
- **Efficiency**: Proportion of correctly identified individuals in each group [33]. If the null hypothesis is that an individual is from the parental species rather than a hybrid individual, then low efficiency suggests a

high rate of type II errors, in which hybrid individuals are incorrectly assigned as parental species.

- **F_{ST}** –A measure of genetic differentiation between populations based on the difference in allele frequencies within and between populations [64]
- **Hybridization:** mating of individuals from diverged populations
- **Hybrid:** an individual that has an intermediate genotype between two diverged, parental populations, as the result of interbreeding between these populations
- **Hybrid swarm:** a hybridizing population that includes F1 hybrids and various backcrosses, due to a total breakdown of assortative mating. Also known as a unimodal hybrid zone [22].
- **Introgression:** the movement of alleles between genetically differentiated forms (including populations, species, etc), mediated by backcrossing [65]
- **Secondary contact:** Occurs when two (or more) species that have been in allopatry come back into sympatry
- **Sympatry**: species in overlapping ranges

Figure 1: Anthropogenic hybridization falls into three main categories. These are 1) systems with inviable or infertile hybrids, 2) bimodal hybrid zones in which there is either mating preference for parental species phenotypes or the relative abundance of parental species means most matings are backcrosses and 3) hybrid swarms in which there is random mating and many hybrid individuals. In this schematic figure we illustrate for each type of anthropogenic hybridization system how many individuals of each admixture (Q) score might be found and typical distributions of mating success across Q scores according to whether there is a high likelihood of hybrid individuals mating with the parental species phenotypes present. While we represent hybrid swarms and bimodal hybrid zones as categorically different, these are probably ends of a continuum and some systems may be intermediate between them. Note that we have represented (2) as a bimodal hybrid zone due to backcrossing into both parental species. Alternatively there can be a single (i.e. unimodal) hybrid zone due to unidirectional backcrossing.

Box 1 – How many markers do I need to discover backcrossed individuals in my system?

Substantial power is needed to detect individuals that are the result of repeated generations of backcrossing. General rules have been suggested, including that for each additional generation twice as many markers are needed [44], and that at least 48 markers would be needed to consistently detect first generation backcrossing in hybrids with parental species that have an F_{st} =0.21 [33]. However, we are now in the age of genomics, when the cost of increasing marker density is dramatically decreasing [58], and thus marker numbers should be less of a barrier than previously. So, how many markers does a study need to reliably detect backcrossed individuals?

To maximize detection of backcrossed individuals, researchers can increase their power in three ways; through increased divergence, the use of diagnostic markers, or with increased numbers of markers. Studies with high divergence between hybridizing species have high power [33]. However, as many conservation biologists choose their study system based on conservation concerns and not to maximize power, this advice is not helpful. Diagnostic markers have fixed allelic differences between parental species and are the most powerful for backcross detection [25]. Ancestry informative markers, those with strong allele frequency divergence between species, are also very powerful [63]. Loci with weak allele frequency divergence between species are least useful. Diagnostic and ancestry informative markers can be determined based on genotyping and contrasting known parental species individuals, although this is not always feasible (e.g. [55]). Additionally, the diagnostic properties of markers are a function of the populations and individuals that have been sampled; more extensive sampling sometimes demonstrates that selected markers are not diagnostic for all populations [66]. Generally speaking, the more markers used, the higher the chance of detection of admixture in an individual [33, 44].

Assuming diagnostic markers, it is ideal to know the number of elapsed generations since the initial hybridization, as, for every further generation of

backcrossing, the proportion of introgressed genome halves [44]. The number of generations since hybridization should be interpreted with an eye to policy. After some number of generations of uni-directional backcrossing, policy will dictate that we consider an individual to be parental species (again) [67]. It's best to make this decision prior to marker selection, as it is impossible to apply policy decisions regarding the acceptability of backcrossed individuals without sufficient detection power.

If we are interested in all generations of backcrossing, then we can extend the deterministic model developed by Boecklen and Howard ([44]; Equation 2) for the genomics era. We made the same assumptions, specifically that backcrossing is unidirectional, loci are independent and Mendelian, all markers are diagnostic, all backcrossing is between the previous generation of backcrosses and parental species, and all genotypes are equally fecund [44]. We asked what proportion of backcrossed individuals are undetectable because they are homozygous for all diagnostic markers. We modeled 10 generations of backcrossing, and each of 10, 100 and 1000 diagnostic markers (Figure I). When using 10 diagnostic markers, 52% of 4^{th} generation backcrosses are homozygous for one parental species at all loci, and thus undetectable as backcrosses. In contrast, 1000 diagnostic markers allow for powerful (85%) detection of 9^{th} generation backcrosses.

Figure I: An extension of the deterministic model presented by Boecklen and Howard [44]. The proportion of hybrid individuals that are homozygous at all the (diagnostic) markers, and are hence indistinguishable from the parental species that is being introgressed, increases with each generation of backcrossing, but decreases with increased marker density. This demonstrates that more markers than are typically used in studies of anthropogenic hybridization are needed to detect advanced backcrosses.

Box 2 – Reporting Error

Credible (or confidence) intervals (CIs) are a powerful, intuitive way to assess confidence in the estimates being presented [68, 69]. Measures of uncertainty are not always presented in estimates of anthropogenic hybridization (although see [53, 70-73] for exceptions), perhaps because the uncertainty is so high where estimated. Credible intervals can be calculated using STRUCTURE [25] and standard errors can be calculated using ADMIXTURE [29], so reporting of error estimates is easily implemented in a routine workflow.

There are practical implications of the reporting of credible intervals, particularly for individuals with very low or very high admixture values (Q). Cut-off thresholds have been used to determine if individuals are members of the parental populations or are admixed, but these thresholds are usually based on the detection power of a study (see main text). Since these are hard cut-offs, individuals with very similar levels of admixture can be assigned to very different populations. For example, with a Q cut-off of 0.80, if individual 1 is assessed as Q=0.79, it is determined to be admixed and, depending on the management of the system, may be culled. In contrast, if individual 2 is estimated to have Q=0.81, it would be considered a parental species individual and be retained for breeding. There may be no substantive difference between these individuals, although this is impossible to tell using only point estimates.

We recommend that credible intervals should also be included in visual depictions of admixture. Typically, the key figure from a paper on anthropogenic hybridization is the characteristic “STRUCTURE Bar Plot” [25], that uses stacked colours to denote genetic contributions from different source populations. These plots show the point estimates for each individual, and allow the author to determine thresholds for inclusion in each group. While such figures are compelling and easily interpreted, they do not convey the uncertainty around individual point estimates.

Allendorf and colleagues [11] noted that it is very difficult to make policy decisions when comparing different low point estimates of admixture. We

573 recommend that researchers should focus on the uncertainty around Q estimates
574 when making decisions about the genetic group each individual belongs to. It has
575 been pointed out that the use of credible intervals demonstrates the high levels
576 of uncertainty researchers are facing [70]. As they should! This problem will of
577 course be substantially alleviated by using more markers (see Box 1).
578

Box 3: Lessons from Natural Systems

Naturally occurring hybrid zones have long been used as 'natural laboratories' to study the speciation process [74]. The field of speciation genomics works to understand how genomic differences build up to cause eventual reproductive isolation [75-78]. Recently, population geneticists have used genome wide markers to ask questions regarding the genomic architecture of reproductive isolation and speciation, and how the genomes of diverged populations change in the face of on-going gene flow [43, 78, 79]. Further, many studies of natural hybrid zones have focused on isolating signals from historical vs. contemporary hybridization (main text 2.1.1, [78]). These questions that speciation biologists ask using hybrid zones could equally be asked in anthropogenic hybrid zones, particularly in studies that used whole genome sequence data. Indeed, studies of anthropogenic hybrid zones may even have more power than those with naturally occurring secondary contact as in some cases of introduced or escaped heterospecifics, phenotypic divergence is more extreme, meaning that fewer individuals would need to be sampled for, for example, admixture mapping [78].

Use of genomic data allows speciation geneticists to examine heterogeneity in divergence across the genome. Indeed, the questions we noted above are most interesting when heterogeneity is found. Genome scans look for regions of high divergence between species (F_{st} or $\mathbf{d_{xy}}$) which may indicate regions that resist introgression, also known as 'speciation islands' [80], or 'islands of differentiation' [79]. While such signals are not without controversy [81], and in some cases may represent phylogenetically derived regions of low recombination, rather than reproductive isolation [82], they represent interesting candidate regions for fixed differences between hybridizing species, and thus could be used diagnostically by conservation biologists. For example, golden-winged (*Vermivora chrysoptera*) and blue-winged warblers (*V. cyanoptera*), which hybridize in eastern North America are phenotypically distinct but undistinguishable when using low density, microsatellite marker panels [83]. Only with the use of whole genome sequencing were six small divergent regions of the genome discovered, four of which are associated with

612 either pigmentation or feather development genes and explain more than 90% of
613 the variation in plumage [39]. This demonstrates that a focus on the use of high
614 density markers to explore heterogeneity across the genome allows for higher
615 power to both distinguish between closely related, hybridizing species
616 genetically, and to associate genomic regions with diverged phenotypes, two
617 possible goals of conservation biologists working on anthropogenic hybrid
618 zones. We echo the call of [1] that conservation biologists can take a cue from
619 speciation biologists that have, in many cases, developed methods that use
620 genomics to ask interesting questions of hybrid zones.
621

References

1. Grabenstein, K.C. and Taylor, S.A. (2018) Breaking Barriers: Causes, Consequences, and Experimental Utility of Human-Mediated Hybridization. Trends in Ecology & Evolution 33(3): 198-212

2. Brennan, A.C. et al. (2015) Hybridization due to changing species distributions: adding problems or solutions to conservation of biodiversity during global change? Evolutionary Ecology Research 16 (6), 475-491.

3. Wayne, R.K. and Shaffer, H.B. (2016) Hybridization and endangered species protection in the molecular era. Molecular Ecology 25 (11), 2680-2689.

4. Crispo, E. et al. (2011) Broken barriers: Human‐induced changes to gene flow and introgression in animals. BioEssays 33 (7), 508-518.

5. Pimm, S.L. et al. (2006) The genetic rescue of the Florida panther. Animal Conservation 9 (2), 115-122.

6. Frankham, R. (2015) Genetic rescue of small inbred populations: Meta‐analysis reveals large and consistent benefits of gene flow. Molecular Ecology 24 (11), 2610-2618.

7. Suarez‐Gonzalez, A. et al. (2016) Genomic and functional approaches reveal a case of adaptive introgression from *Populus balsamifera* (balsam poplar) in *P. átrichocarpa* (black cottonwood). Molecular Ecology 25 (11), 2427-2442.

8. Hamilton, J.A. and Miller, J.M. (2016) Adaptive introgression as a resource for management and genetic conservation in a changing climate. Conservation Biology 30 (1), 33-41.

9. Kovach, R.P. et al. (2016) Risk and efficacy of human‐enabled interspecific hybridization for climate‐change adaptation: response to Hamilton and Miller (2016). Conservation Biology 30 (2), 428-430.

10. Kardos, M. and Shafer, A.B. (2018) The Peril of Gene-Targeted Conservation. Trends in Ecology & Evolution 33(11): 827-839

11. Allendorf, F.W. et al. (2001) The problems with hybrids: setting conservation guidelines. Trends in Ecology & Evolution 16 (11), 613-622.

12. Allendorf, F.W. and Luikart, G. (2009) Conservation and the genetics of populations, John Wiley & Sons.

13. Todesco, M. et al. (2016) Hybridization and extinction. Evolutionary Applications 9 (7), 892-908.

14. Rhymer, J.M. and Simberloff, D. (1996) Extinction by hybridization and introgression. Annual Review of Ecology and Systematics, 83-109.

15. Rieseberg, L.H. et al. (1989) Hybridization in the island endemic, *Catalina mahogany*. Conservation Biology 3 (1), 52-58.

16. Petit, R.J. (2004) Biological invasions at the gene level. Diversity and Distributions 10 (3), 159-165.

17. Schumer, M. et al. (2018) Natural selection interacts with recombination to shape the evolution of hybrid genomes. Science DOI: 10.1126/science.aar3684

18. Fitzpatrick, B.M. et al. (2010) Rapid spread of invasive genes into a threatened native species. Proceedings of the National Academy of Sciences 107 (8), 3606-3610.

19. Amish, S.J. et al. (2012) RAD sequencing yields a high success rate for westslope cutthroat and rainbow trout species‐diagnostic SNP assays. Molecular Ecology Resources 12 (4), 653-660.

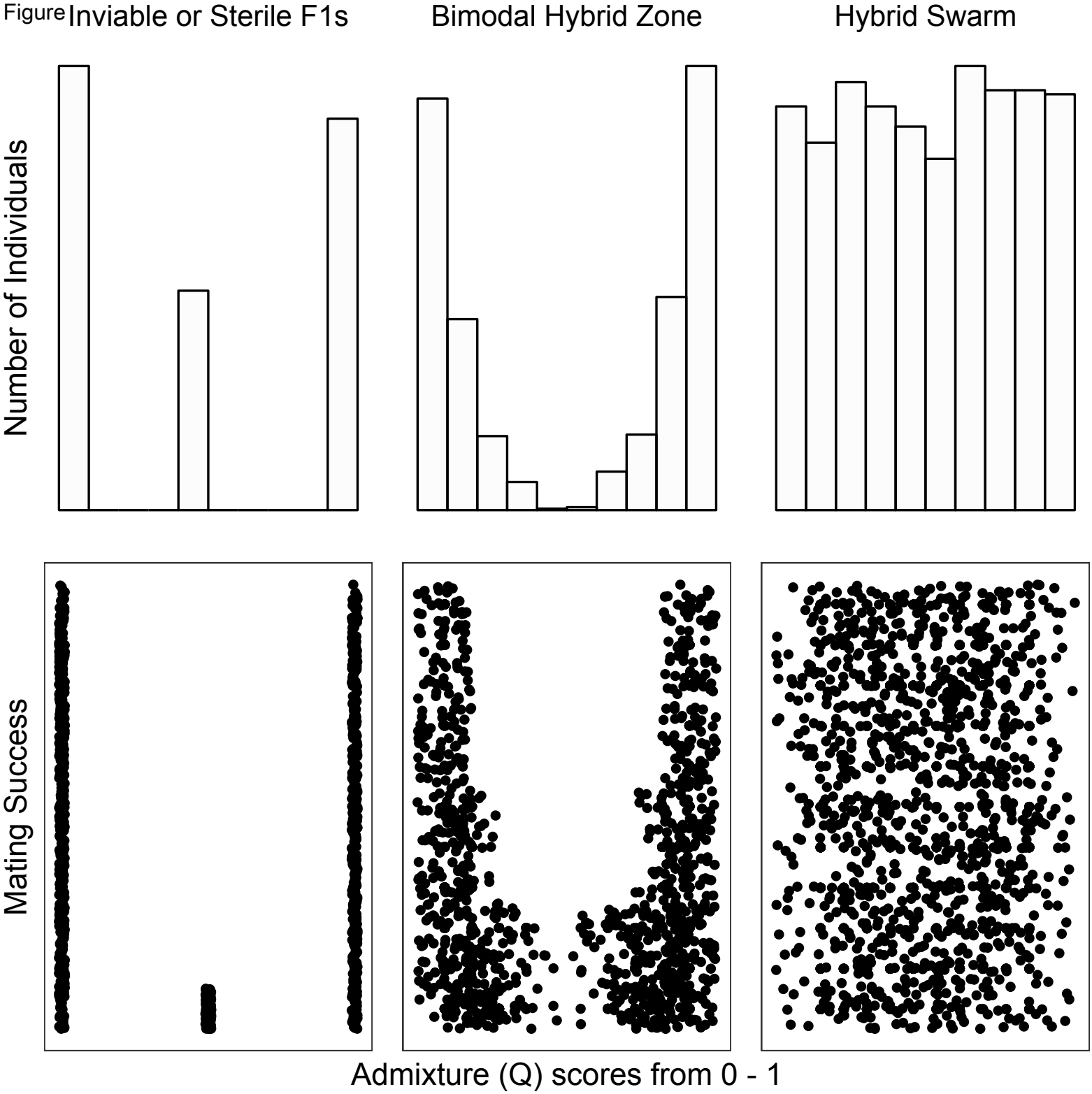
669 20. Epifanio, J. and Philipp, D. (2000) Simulating the extinction of parental
670 lineages from introgressive hybridization: the effects of fitness, initial
671 proportions of parental taxa, and mate choice. Reviews in Fish Biology and
672 Fisheries 10 (3), 339-354.
673 21. Lamichhaney, S. et al. (2018) Rapid hybrid speciation in Darwin's finches.
674 Science 359 (6372), 224-228.
675 22. Jiggins, C.D. and Mallet, J. (2000) Bimodal hybrid zones and speciation.
676 Trends in Ecology & Evolution 15 (6), 250-255.
677 23. Rundle, H.D. and Nosil, P. (2005) Ecological speciation. Ecology Letters 8 (3),
678 336-352.
679 24. Allen, B. et al. (2016) Role of genetic background in the introgressive
680 hybridization of rainbow trout (*Oncorhynchus mykiss*) with Westslope cutthroat
681 trout (*O. clarkii lewisi*). Conservation genetics 17 (3), 521-531.
682 25. Pritchard, J.K. et al. (2000) Inference of population structure using multilocus
683 genotype data. Genetics 155 (2), 945-959.
684 26. Raj, A. et al. (2014) fastSTRUCTURE: variational inference of population
685 structure in large SNP data sets. Genetics 197 (2), 573-589.
686 27. Kwan, Y.S. et al. (2014) Genomic replacement of native *Cobitis lutheri* with
687 introduced *C. tetralineata* through a hybrid swarm following the artificial
688 connection of river systems. Ecology and Evolution 4 (8), 1451-1465.
689 28. Le Roux, J.J. et al. (2015) Genetic analysis shows low levels of hybridization
690 between African wildcats (*Felis silvestris lybica*) and domestic cats (*F. s. catus*) in
691 South Africa. Ecology and Evolution 5 (2), 288-299.
692 29. Alexander, D.H. and Lange, K. (2011) Enhancements to the ADMIXTURE
693 algorithm for individual ancestry estimation. BMC bioinformatics 12 (1), 246.
694 30. Galaverni, M. et al. (2017) Disentangling timing of admixture, patterns of
695 introgression and phenotypic indicators in a hybridizing wolf population.
696 Molecular Biology and Evolution 34(9):2324-2339
697 31. Nielsen, E.E. et al. (2006) HYBRIDLAB (version 1.0): a program for generating
698 simulated hybrids from population samples. Molecular Ecology Resources 6 (4),
699 971-973.
700 32. van Heugten, R.A. et al. (2017) Sleeping with the 'enemy': hybridization of an
701 endangered tree weta. Conservation Genetics 18 (6), 1377-1387.
702 33. Vähä, J.-P. and Primmer, C.R. (2006) Efficiency of model‐based Bayesian
703 methods for detecting hybrid individuals under different hybridization scenarios
704 and with different numbers of loci. Molecular Ecology 15 (1), 63-72.
705 34. vonHolt, B.M. et al. (2016) Whole-genome sequence analysis shows that two
706 endemic species of North American wolf are admixtures of the coyote and gray
707 wolf. Science Advances 2 (7), e1501714.
708 35. Hohenlohe, P.A. et al. (2017) Comment on "Whole-genome sequence analysis
709 shows two endemic species of North American wolf are admixtures of the coyote
710 and gray wolf". Science Advances 3 (6), e1602250.
711 36. Durand, E.Y. et al. (2011) Testing for ancient admixture between closely
712 related populations. Molecular Biology and Evolution 28 (8), 2239-2252.
713 37. Soraggi, S. et al. (2018) Powerful Inference with the D-Statistic on Low-
714 Coverage Whole-Genome Data. G3: Genes, Genomes, Genetics 8 (2), 551-566.
715 38. Gutenkunst, R.N. et al. (2009) Inferring the joint demographic history of
716 multiple populations from multidimensional SNP frequency data. PLoS Genetics
717 5 (10), e1000695.

718 39. Toews, D.P. et al. (2016) Plumage genes and little else distinguish the
719 genomes of hybridizing warblers. Current Biology 26 (17), 2313-2318.
720 40. Gravel, S. (2012) Population genetics models of local ancestry. Genetics
721 191(2):607-619
722 41. Harris, K. and Nielsen, R. (2013) Inferring demographic history from a
723 spectrum of shared haplotype lengths. PLoS Genetics 9 (6), e1003521.
724 42. Palamara, P.F. et al. (2012) Length distributions of identity by descent reveal
725 fine-scale demographic history. The American Journal of Human Genetics 91 (5),
726 809-822.
727 43. Payseur, B.A. and Rieseberg, L.H. (2016) A genomic perspective on
728 hybridization and speciation. Molecular Ecology 25 (11), 2337-2360.
729 44. Boecklen, W.J. and Howard, D.J. (1997) Genetic analysis of hybrid zones:
730 numbers of markers and power of resolution. Ecology 78 (8), 2611-2616.
731 45. Dumont, B.L. (2017) Variation and evolution of the meiotic requirement for
732 crossing over in mammals. Genetics 205 (1), 155-168.
733 46. Johnston, S.E. et al. (2018) A genomic region containing REC8 and RNF212B
734 is associated with individual recombination rate variation in a wild population of
735 red deer (*Cervus elaphus*). 8(7):2265-2276
736 47. Ellegren, H. (2010) Evolutionary stasis: the stable chromosomes of birds.
737 Trends in Ecology & Evolution 25 (5), 283-291.
738 48. Anderson, E.C. (2009) Statistical methods for identifying hybrids and groups.
739 In Population Genetics and Animal Conservation (Bertorelle, G. et al. eds), pp. 25-
740 41, Cambridge University Press.
741 49. Randi, E. (2008) Detecting hybridization between wild species and their
742 domesticated relatives. Molecular Ecology 17 (1), 285-293.
743 50. Galov, A. et al. (2015) First evidence of hybridization between golden jackal
744 (*Canis aureus*) and domestic dog (*Canis familiaris*) as revealed by genetic
745 markers. Royal Society Open Science 2 (12), 150450.
746 51. Haldane, J.B. (1922) Sex ratio and unisexual sterility in hybrid animals.
747 Journal of Genetics 12 (2), 101-109.
748 52. Turelli, M. and Moyle, L.C. (2007) Asymmetric postmating isolation: Darwin's
749 corollary to Haldane's rule. Genetics 176 (2), 1059-1088.
750 53. Senn, H.V. and Pemberton, J.M. (2009) Variable extent of hybridization
751 between invasive sika (*Cervus nippon*) and native red deer (*C. elaphus*) in a small
752 geographical area. Molecular Ecology 18 (5), 862-876.
753 54. Schulte, U. et al. (2012) Rapid genetic assimilation of native wall lizard
754 populations (*Podarcis muralis*) through extensive hybridization with introduced
755 lineages. Molecular Ecology 21 (17), 4313-4326.
756 55. Pilot, M. et al. (2018) Widespread, long‐term admixture between grey
757 wolves and domestic dogs across Eurasia and its implications for the
758 conservation status of hybrids. Evolutionary Applications.
759 DOI:10.1111/eva.12595
760 56. Hohenlohe, P.A. et al. (2013) Genomic patterns of introgression in rainbow
761 and westslope cutthroat trout illuminated by overlapping paired‐end RAD
762 sequencing. Molecular Ecology 22 (11), 3002-3013.
763 57. Boyer, M.C. et al. (2008) Rainbow trout (*Oncorhynchus mykiss*) invasion and
764 the spread of hybridization with native westslope cutthroat trout (*Oncorhynchus*
765 *clarkii lewisi*). Canadian Journal of Fisheries and Aquatic Sciences 65 (4), 658-
766 669.

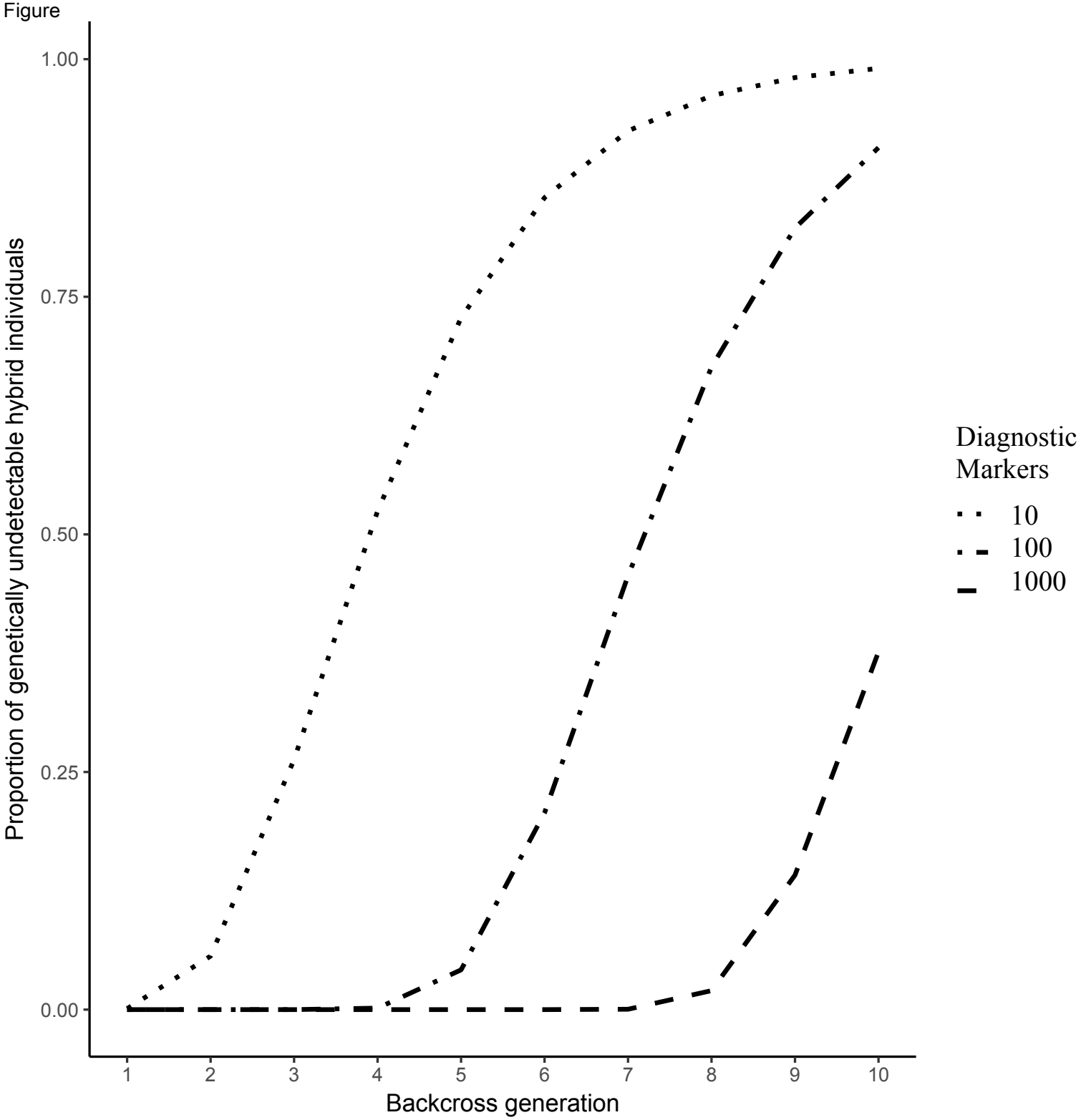
767 58. Ellegren, H. (2014) Genome sequencing and population genomics in non-
768 model organisms. Trends in Ecology & Evolution 29 (1), 51-63.
769 59. van der Valk, T. et al. (2017) Whole mitochondrial genome capture from
770 faecal samples and museum‐preserved specimens. Molecular Ecology
771 Resources 17 (6), e111-e121.
772 60. de Groot, G.A. et al. (2016) Decades of population genetic research reveal the
773 need for harmonization of molecular markers: the grey wolf *Canis lupus* as a case
774 study. Mammal Review 46 (1), 44-59.
775 61. Andrews, K.R. et al. (2016) Harnessing the power of RADseq for ecological
776 and evolutionary genomics. Nature Reviews Genetics 17 (2), 81-92.
777 62. Koskinen, M.T. et al. (2004) The benefits of increasing the number of
778 microsatellites utilized in genetic population studies: an empirical perspective.
779 Hereditas 141 (1), 61-67.
780 63. Shriver, M.D. et al. (2003) Skin pigmentation, biogeographical ancestry and
781 admixture mapping. Human Genetics 112 (4), 387-399.
782 64. Wright, S. (1943) Isolation by distance. Genetics 28, 114-138.
783 65. Avise, J.C. (2012) Molecular markers, natural history and evolution, Springer
784 Science & Business Media.
785 66. Hand, B.K. et al. (2015) Genomics and introgression: discovery and mapping
786 of thousands of species-diagnostic SNPs using RAD sequencing. Current Zoology
787 61 (1), 146-154.
788 67. Lavretsky, P. et al. (2016) Becoming pure: identifying generational classes of
789 admixed individuals within lesser and greater scaup populations. Molecular
790 Ecology 25 (3), 661-674.
791 68. Altman, D. et al. (2013) Statistics with confidence: confidence intervals and
792 statistical guidelines, John Wiley & Sons.
793 69. Nakagawa, S. and Cuthill, I.C. (2007) Effect size, confidence interval and
794 statistical significance: a practical guide for biologists. Biological Reviews 82 (4),
795 591-605.
796 70. Bohling, J.H. and Waits, L.P. (2011) Assessing the prevalence of hybridization
797 between sympatric Canis species surrounding the red wolf (*Canis rufus*)
798 recovery area in North Carolina. Molecular Ecology 20 (10), 2142-2156.
799 71. Trigo, T. et al. (2008) Inter‐species hybridization among Neotropical cats of
800 the genus *Leopardus*, and evidence for an introgressive hybrid zone between *L.*
801 *geoffroyi* and *L. tigrinus* in southern Brazil. Molecular Ecology 17 (19), 4317-
802 4333.
803 72. Beaumont, M. et al. (2001) Genetic diversity and introgression in the Scottish
804 wildcat. Molecular Ecology 10 (2), 319-336.
805 73. Yokoyama, R. et al. (2009) Disturbance of the indigenous gene pool of the
806 threatened brook lamprey *Lethenteron* sp. S by intraspecific introgression and
807 habitat fragmentation. Conservation genetics 10 (1), 29-43.
808 74. Hewitt, G.M. (1988) Hybrid zones-natural laboratories for evolutionary
809 studies. Trends in Ecology & Evolution 3 (7), 158-167.
810 75. Nosil, P. and Feder, J. (2012) Genomic divergence during speciation, causes
811 and consequences. Philosophical Transations of the Royal Society B 367(1587):
812 76. Seehausen, O. et al. (2014) Genomics and the origin of species. Nature
813 Reviews Genetics 15 (3), 176.
814 77. Butlin, R.K. (2010) Population genomics and speciation. Genetica 138 (4),
815 409-418.

816 78. Nadeau, N.J. and Kawakami, T. (2018) Population Genomics of Speciation and
817 Admixture. *in* Population Genomics. Springer, Cham.
818 DOI:10.1007/13836_2018_24
819 79. Wolf, J.B. and Ellegren, H. (2017) Making sense of genomic islands of
820 differentiation in light of speciation. Nature Reviews Genetics 18 (2), 87.
821 80. Turner, T.L. et al. (2005) Genomic islands of speciation in *Anopheles gambiae*.
822 PLoS Biology 3 (9), e285.
823 81. Cruickshank, T.E. and Hahn, M.W. (2014) Reanalysis suggests that genomic
824 islands of speciation are due to reduced diversity, not reduced gene flow.
825 Molecular Ecology 23 (13), 3133-3157.
826 82. Burri, R. et al. (2015) Linked selection and recombination rate variation drive
827 the evolution of the genomic landscape of differentiation across the speciation
828 continuum of *Ficedula* flycatchers. Genome Research 25 (11), 1656-1665.
829 83. Vallender, R. et al. (2007) Complex hybridization dynamics between
830 golden‐winged and blue‐winged warblers (*Vermivora chrysoptera* and
831 *Vermivora pinus*) revealed by AFLP, microsatellite, intron and mtDNA markers.
832 Molecular Ecology 16 (10), 2017-2029.

Figure


Inviable or Sterile F1s
Bimodal Hybrid Zone
Hybrid Swarm
Number of Individuals
Mating Success
Admixture (Q) scores from 0 - 1

Figure


1.00
0.75
0.50
0.25
0.00
Proportion of genetically undetectable hybrid individuals
1
2
3
4
5
6
7
8
9
10
Backcross generation
Diagnostic Markers
10
100
1000

Outstanding questions:

1) Do replicate anthropogenic hybrid zones show similar patterns of introgression?

There are big evolutionary questions that could be answered by the sorts of data that conservation biologists working on anthropogenic hybridization could answer. For example, there are multiple replicate hybrid zones occurring in the wolf/dog, wild cat/domestic cat, red deer/sika deer, westslope cut throat trout/rainbow trout systems. But in many cases, there is limited communication and collaboration between researchers, or different markers are used across studies [60]. Clearly this isn't a problem unique to this field, but it is the case that collaboration between researchers would be made easier with standardized genome wide data aligned to a common genome. Genomic data make cross study comparisons easier, and would allow for easier comparison between studies.

2) Once there has been a breakdown of reproductive isolation characterized as hybridization, how common is maintenance of within parental species assortative mating? Is the strength of assortative mating stronger when species are more diverged, or perhaps between closely related species that have recently evolved reproductive isolation?

3) What is the relative frequency of hybrid swarms vs bimodal hybrid zones? We expect that the prevalence of bimodal hybrid zones has been underestimated because of the difficulty of detecting highly introgressed backcrosses. Increased use of high-density markers will make these cases easier to detect and would enhance our understanding of the systems that are bimodal hybrid zones.

Highlights:

Anthropogenic hybridization is increasingly common and likely to result in a breakdown of reproductive isolation between 'good' species.

Backcrossed individuals that have only a small proportion of one parental genome are difficult to differentiate from parental individuals using the most common current technologies.

Bimodal hybrid zones are characterized by introgression and backcrossing. The majority of hybrid individuals in these systems have low levels of introgression. The problems posed by bimodal hybrid zones have been largely overlooked in the literature.

Genome wide sampling of genetic markers at high densities allow for increased precision in the estimate of admixture proportions, which makes it feasible to detect multi-generation backcrosses, and will thus make it easier to differentiate bimodal hybrid zones from hybrid swarms or systems without introgression.